CHALMERS
UNIVERSITY OF TECHNOLOGY

Analyzing Factors Influencing Pedestrian Behavior in Urban Traffic Scenarios using Deep Learning

Citation for the original published paper (version of record):
Zhang, C., Berger, C. (2023). Analyzing Factors Influencing Pedestrian Behavior in Urban Traffic Scenarios using Deep Learning. Transportation Research Procedia, 72.
http://dx.doi.org/10.1016/j.trpro.2023.11.637

(article starts on next page)

Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 72 (2023) 1653–1660

Transportation Research Procedia

www.elsevier.com/locate/procedia

Transport Research Arena (TRA) Conference

Analyzing Factors Influencing Pedestrian Behavior in Urban Traffic Scenarios using Deep Learning

Chi Zhang [a,*], Christian Berger [a]

[a]*Department of Computer Science and Engineering, University of Gothenburg, 412 96 Gothenburg, Sweden*

Abstract

Pedestrians are the most vulnerable road users in urban traffic scenarios and need to be protected from potentially hazardous situations. It is essential for automated vehicles and modern driver-assistance systems to better predict pedestrians' behavior to prevent road crashes. Predicting pedestrian behavior is challenging because their behavior can be influenced by many factors. In recent years, deep learning (DL) methods as powerful tools have been utilized by many researchers to improve such predictions, but few researchers have analyzed the factors that influence pedestrian behavior prediction in DL. This paper uses DL to predict and analyze the factors that influence pedestrian behavior, especially the interactions between pedestrians and other road users. We focus on real-world urban traffic and use the publicly available Waymo Open Dataset for training, testing, and analyzing.

Keywords: Deep learning; pedestrian behavior prediction; trajectory prediction; pedestrian interactions; behavior analysis; automated vehicles.

1. Introduction

According to WHO's report on road safety WHO (2018), more than 310,000 pedestrians lose their lives because of road crashes every year, constituting 23% of all road deaths. The unacceptably high fatalities show a great demand for reducing such accidents on the road, especially for protecting the pedestrians who are the most vulnerable road users. Therefore, understanding and predicting pedestrian behavior is crucial because it can contribute to preventing hazardous situations.

However, predicting pedestrian behavior is a very challenging task because pedestrians can change their speeds and directions unexpectedly (cf. Schneider and Gavrila (2013)), and they tend to interact with the other road users all the time (cf. Ohn-Bar and Trivedi (2016), Shirazi and Morris (2015), Rasouli and Tsotsos (2019)). The randomness and

* Corresponding author. Tel.: +46-(0)72-974-5016.
E-mail address: chi.zhang@gu.se

10.1016/j.trpro.2023.11.637

interactions lead to the complexity and non-linearity of pedestrian behavior prediction, and therefore, it is hard to use rule-based and statistics-based methods for prediction. In recent years, deep learning (DL) methods that can deal with non-linear problems have been utilized by many researchers for predicting pedestrian behavior, such as the trajectory prediction as proposed by Nikhil and Morris (2018), Giuliari et al. (2021), Alahi et al. (2016), Gupta et al. (2018), Sadeghian et al. (2019), and Song et al. (2020), and the intention prediction as proposed by Zhang et al. (2020), Yang et al. (2021), and Kotseruba et al. (2021). In this paper, we mainly focus on pedestrian trajectory prediction.

Many research studies focused on utilizing various factors of interactions in pedestrian behavior prediction, such as social interaction between pedestrians proposed by Alahi et al. (2016), interactions between pedestrians and all vehicles on the road proposed by Zhang and Berger (2022), and interactions between pedestrians and a single vehicle proposed by Eiffert et al. (2020), which could be either an automated vehicle or a manually controlled vehicle. However, there is not much work looking into the influence on prediction accuracy of using these factors in deep learning models. This paper analyzes the factors that influence pedestrian behavior prediction in deep learning models, especially the impact of interaction factors on trajectory prediction. By analyzing existing influencing factors, we select the factors that influence the trajectory prediction the most, and use them to improve prediction accuracy.

Our main contributions are as follows:

- We propose a framework for pedestrian trajectory prediction, and analyze the factors that influence pedestrian behavior in deep learning models. In addition to social interaction and pedestrian-vehicle interaction used by models proposed previously, we also study the interaction between pedestrians and the automated ego-vehicle mounted with sensors.
- We investigate both sequential and non-sequential deep learning networks, and study how the factors can influence the prediction of a pedestrian's behavior.
- We use a real-world dataset (Waymo Open Dataset proposed by Sun et al. (2020)) for predicting and analyzing, which is complicated containing multiple road users and various factors that influence the behavior of pedestrians.

2. Related Works

2.1. Factors that Influence Pedestrian Trajectory Prediction

Various factors can influence pedestrian trajectory and are considered in existing deep learning models. Previous existing deep learning models focused primarily on the following:

- Target pedestrians. The factors related to the target pedestrians themselves, including past trajectories as used by Nikhil and Morris (2018) and Giuliari et al. (2021), and moving states such as velocity or orientation as used by Song et al. (2020).
- Interaction with other road users. The factors related to interactions with other road users, including the social interaction with other pedestrians as used by Alahi et al. (2016), Gupta et al. (2018), Sadeghian et al. (2019), Mohamed et al. (2020), and Zhang et al. (2021), interactions between pedestrians and vehicles as used by Yang et al. (2021), Kotseruba et al. (2021), and Zhang and Berger (2022), and interactions between pedestrians and a single vehicle as proposed by Eiffert et al. (2020), which could be used for modeling the interaction between pedestrians and automated ego-vehicle.
- Environment factors, as used by Sadeghian et al. (2019), Zhang et al. (2020), Yang et al. (2021), and Kotseruba et al. (2021).

This paper mainly focuses on analyzing the factors related to the interaction with other road users. However, most of the previous existing deep learning models for pedestrian trajectory prediction uses datasets that do not include vehicles (e.g., ETH proposed by Pellegrini et al. (2009) and UCY proposed by Lerner et al. (2007)), or the datasets in use contain limited scenes in a few fixed areas (e.g., Usyd proposed by Zhou et al. (2020)). These datasets either do not include vehicles or do not contain diverse driving scenarios on the road. Therefore, to analyze interactions between pedestrians and other road users, we use the Waymo Open Dataset proposed by Sun et al. (2020), which has been collected in real traffic scenarios containing not only pedestrians, but also vehicles and the ego vehicle.

2.2. Deep Learning Models for Trajectory Prediction

Influencing factors can have different impacts on the prediction accuracy when using various deep learning prediction backbones for prediction as discussed by Zhang and Berger (2022). Therefore, we analyze two kinds of prediction backbones, including both, sequential (long short-term memory, LSTM-based models) and non-sequential (convolutional, Conv-based models) in this paper. Here we briefly introduce these two prediction backbones.

LSTM-based methods: LSTM networks are an improved version of recurrent neural networks (RNNs) that have both, feedforward and feedback connections, and can capture both long and short-term temporal information. LSTMs are especially suitable for pedestrian trajectory prediction, which is based on sequential data, and have been utilized by many researchers (cf. Alahi et al. (2016), Fernando et al. (2018), Zhang et al. (2019), Zhang and Berger (2022)).

Conv-based methods: The convolutional networks including convolutional neural networks (CNNs) and temporal convolutional networks (TCNs), are the networks that use convolutional layers for extracting the temporal feature and can be used for predicting trajectories. Nikhil and Morris (2018) utilized CNNs to predict trajectories, reaching competitive results with a faster inference speed. Mohamed et al. (2020), Zhang et al. (2021), and Zhang and Berger (2022) used TCNs and CNNs for trajectory prediction and achieved a better performance than LSTM-based models.

3. Methodology

3.1. Problem Definition

In this paper, we use observed trajectories of pedestrians and vehicles in the past to predict pedestrian trajectories in the future. The illustration of the problem is shown in Fig. 1. In a frame at time-step t with the number of pedestrians n_p and the number of vehicles n_v, we define the position of i^{th} person as $\boldsymbol{X}_t^i=(x_t^i, y_t^i)$, where $i \in \{1, \ldots, n_p\}$, and the position of the j^{th} vehicle as $\boldsymbol{V}_t^j=(x_t^j, y_t^j)$, where $j \in \{1, \ldots, n_v\}$. The trajectories of observed pedestrians and vehicles can be defined sequences of x-y coordinate positions, and be denoted as $\boldsymbol{X}_t=\left[X_t^1, X_t^2, \ldots, X_t^{n_p}\right]$, $\boldsymbol{V}_t=\left[V_t^1, V_t^2, \ldots, V_t^{n_v}\right]$ with all observed time-steps $1 \leq t \leq T_{obs}$.

Given the observed pedestrians and vehicles, we aim to predict the most likely trajectories of pedestrians $\widehat{\boldsymbol{Y}}_t=\left[\widehat{Y}_t^1, \widehat{Y}_t^2, \ldots, \widehat{Y}_t^{n_p}\right]$ in the future time-steps $T_{obs} + 1 \leq t \leq T_{pred}$. The ground truth of future trajectories is denoted as $\boldsymbol{Y}_t=\left[Y_t^1, Y_t^2, \ldots, Y_t^{n_p}\right]$, where $T_{obs} + 1 \leq t \leq T_{pred}$.

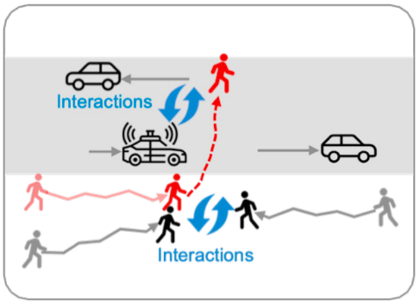

Fig. 1. The illustration of pedestrian trajectory prediction. We consider the interactions with the other pedestrians, vehicles, and the automated vehicle for each pedestrian.

3.2. Influencing Factors Used for Prediction

The overall prediction framework is shown in Fig. 2. In this paper, we include the following factors when predicting a pedestrian's trajectory:

1. Past trajectories of pedestrians. This is the factor related to the target pedestrians themselves. The LSTM model and Conv model without any interactions are compared as baselines.

2. Social interaction (SI) with other pedestrians. This is the factor related to the interaction with other pedestrians. For sequential models, Social-LSTM proposed by Alahi et al. (2016) and SI-LSTM proposed by Zhang and Berger (2022) are compared. For non-sequential models, Social-IWSTCNN proposed by Zhang et al. (2021) is compared with the baseline.
3. Pedestrian-vehicle interaction (PVI). This is the factor related to the interaction between pedestrians and vehicles. For sequential models, PVI-LSTM and SI-PVI-LSTM proposed by Zhang and Berger (2022) are compared. For non-sequential models, PVI-Conv and SI-PVI-Conv proposed by Zhang and Berger (2022) are compared.
4. Pedestrian and ego-vehicle interaction (PEI). This is the factor related to the interaction between pedestrians and an automated vehicle. The automated vehicle is mounted with sensors, and can also be called the ego-vehicle as we are interested in its perception. In this paper, we consider the PEI feature and proposed the following models. For sequential models, we propose: a) the PEI-LSTM that considers the pedestrian and ego-vehicle interaction feature, and b) the SI-PEI-LSTM that also considers the social interaction feature, and c) the SI-PVI-PEI-LSTM that includes all three kinds of interaction features are compared. For non-sequential models, we propose a) the PEI-Conv, b) SI-PEI-Conv, and c) SI-PVI-PEI-Conv model. These proposed models are compared with the other baseline models to analyze the influence of pedestrian and ego-vehicle on pedestrian behavior prediction.

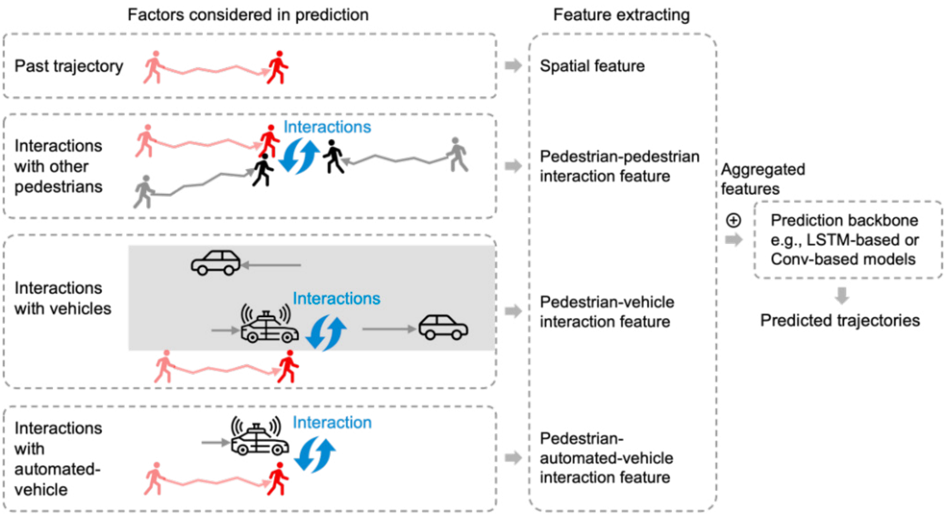

Fig. 2. The overall framework of pedestrian trajectory prediction.

For the extraction of the pedestrian and ego-vehicle interaction feature, we use relative positions, relative velocities, and distance between pedestrians and the ego-vehicle as the input for feature extraction. Linear embedding followed by multi-layer perceptrons (MLPs) are used to extract the interaction relationship between pedestrians and the ego-vehicle. The interaction relationship is aggregated with the ego-vehicle state extracted by the encoder to obtain the pedestrian and ego-vehicle interaction feature.

3.3. Evaluation Metrics for Analyzing

To compare and analyze the prediction accuracy, we use the following metrics to evaluate the prediction error:

- The Average Displacement Error (ADE): The average distance between ground truth and prediction trajectories over all predicted time-steps as defined:

$$ADE = \frac{\sum_{i \in n_p} \sum_{t=T_{obs}+1}^{T_{pred}} \left\| Y_t^i - \widehat{Y}_t^i \right\|_2}{n_p \times \left(T_{pred} - T_{obs} \right)} \tag{1}$$

- The Final Displacement Error (FDE): The average distance between ground truth and prediction trajectories for the final predicted time-step as defined below, where n_p is the number of pedestrians.

$$FDE = \frac{\sum_{i \in n_p} \left\| Y_t^i - \widehat{Y}_t^i \right\|_2}{n_p}, \; t = T_{pred} \tag{2}$$

3.4. Dataset Introduction and Data Pre-Processing

When predicting pedestrian trajectories in urban traffic scenarios, there are drawbacks of the datasets that are commonly used previously (e.g. ETH and UCY datasets) as we discussed in Sec. 2.1. In this paper, we use the Waymo Open Dataset proposed by Sun et al. (2020) that was collected in real-world traffic to analyze the pedestrian trajectories. There are 374 training records and 76 test records collected in urban traffic. The records are collected by a top-mounted LiDAR sensor with 75 m scan range. Pedestrians and vehicles are labeled by 3D bounding boxes with center position (*x, y, z*) and size. In this paper, we consider the pedestrians and vehicles as points regardless of their sizes, and use 2D map representation (*x, y*) from the bird's-eye-view.

We randomly separated the training dataset into 90% (337 scenarios) for training and 10% (37 scenarios) for validation. To align the prediction with the previous works proposed by Zhang et al. (2021), Zhang and Berger (2022) and with most of the trajectory prediction models, we down-sampled the sequences from 10 fps to 2.5 fps. The sequences are divided into pieces with a length of 20 time-steps (i.e., 8 seconds), where the first 8 time-steps (i.e., 3.2 seconds) are used for observation and the last 12 time-steps (i.e., 4.8 seconds) are used for prediction and evaluation. To augment the data amount, there are overlaps between the divided sequences. The skip step is set as one time-step. The data amount for training and testing is listed in Table 1.

Table 1. The number of scenarios and sequences of the Waymo Open Dataset used for training, validation, and evaluation

	Training	Validation	Evaluation	Total
Number of scenarios	337	37	76	450
Number of sequences of scenes	7337	991	1978	10,306
Number of sequences of pedestrians	195,192	36,946	52,484	284,622

The data is recorded from the vehicle's view with the center of ego-vehicle as the origin of each frame in local coordinates. Using the local coordinate will affect the accuracy of prediction because it will introduce the movement of the ego-vehicle into the pedestrian's movement. To reduce the effect of the ego-vehicle's movement, we transform the coordinates from local to global, using the ego-vehicle's start position of each scenario as the origin.

In this paper, we investigate the interaction between pedestrians and ego-vehicle. However, not all pedestrians interact with the ego-vehicle. For example, the pedestrians behind the ego-vehicle are not likely to interact with it. Besides, we mainly pay attention to the pedestrians in front while driving. Therefore, we remove pedestrians behind the ego-vehicle, and use pedestrians in front of the ego-vehicle for training and testing.

4. Results and Discussions

In this section, we present and discuss the prediction results of sequential models and non-sequential models. The factors considered in models are listed in tables below, and results for a) all pedestrians in frames, and b) the pedestrians in front of the ego-vehicle with the sensors mounted are presented.

4.1. Sequential Models

For the sequential models, we use LSTM-based models as the prediction backbone. The trajectory prediction results of LSTM-based models are shown in Table 2.

Table 2. The ADE/FDE metrics (in meters) for sequential methods (LSTM-based models) consider different factors on the Waymo Open Dataset: The lower, the better. Traj stands for past trajectories of pedestrians, SI stands for social interactions between pedestrians, PVI stands for pedestrian-vehicle interactions, and PEI stands for pedestrian and ego-vehicle interaction.

Model	Factors considered for prediction				Waymo (all pedestrians)	Waymo (front pedestrians)
	Traj	SI	PVI	PEI	ADE / FDE	ADE / FDE
LSTM	✓				0.392 / 0.844	0.414 / 0.902
Social-LSTM (2016)	✓	✓			0.402 / 0.840	0.440 / 0.942
SI-LSTM (2022)	✓	✓			0.384 / 0.820	**0.396 / 0.848**
PVI-LSTM (2022)	✓		✓		0.372 / 0.798	0.400 / 0.860
SI-PVI-LSTM (2022)	✓	✓	✓		**0.372 / 0.796**	0.414 / 0.886
PEI-LSTM (ours)	✓			✓	-	0.400 / 0.858
SI-PEI-LSTM (ours)	✓	✓		✓	-	0.403 / 0.866
SI-PVI-PEI-LSTM (ours)	✓	✓	✓	✓	-	0.404 / 0.870

We noticed that the ADE and FDE on the pedestrians in front of the ego-vehicle are worse compared to all pedestrians in Waymo Open Dataset. This indicates that the pedestrians in front of the ego-vehicle are more likely to have complicated interactions, and their behaviors are more challenging to predict. For data with all pedestrians, the best performance of ADE and FDE are 0.372 m and 0.796 m by the SI-PVI-LSTM model that used past trajectories, social interaction between pedestrians, and pedestrian-vehicle interaction. For data with the pedestrians in front of the ego-vehicle, the best performance of ADE and FDE are 0.396 m and 0.848 m by SI-LSTM model that used past trajectories and social interactions between pedestrians. This shows that on different types of datasets, the model that gets the best performance could be different. For the data with all pedestrians, they tend to interact with both other pedestrians and the vehicles on the road. But for the pedestrians in front of the ego-vehicle, they could be far from some of the vehicles in the frame. Considering the influence of these vehicles on pedestrians might involve noises and leads to worse results.

From the results of Social-LSTM proposed by Alahi et al. (2016) and SI-LSTM proposed by Zhang and Berger (2022), we find that including the same factors but with different network structures can lead to opposite performance. The Social-LSTM model used here is the improved version proposed by Gupta et al. (2018). Both Social-LSTM and SI-LSTM considered social interactions between pedestrians, but the Social-LSTM did not improve the prediction accuracy compared to the LSTM model, which is consistent with the results by Gupta et al. (2018) and Zhang and Berger (2022). The Social-LSTM used hidden states of LSTMs to represent the motion feature of pedestrians and then extracted the social interaction feature with the pooling sub-network, while SI-LSTM firstly learned the interaction and spatial features of each frame, and then used LSTMs to obtain the social interaction state features. Therefore, the network structures need to be carefully designed to get more accurate predictions in deep learning models.

Compared with the LSTM model, the ADE and FDE are improved by including social interaction features, pedestrian-vehicle interaction features, and pedestrian and ego-vehicle interaction features. However, the combination of these factors has not reduced the prediction error. There are two possible reasons for this: Firstly, the current aggregation method may not be proper to aggregate these factors, as we simply concatenated the extracted features. Secondly, there might be dependencies between these features, which means that including one feature can represent the others. More research needs to be done to investigate a way to better combine these features and to improve the prediction accuracy.

4.2. Non-Sequential Models

For the non-sequential models, we use Conv-based models that embody TCNs and CNNs as prediction backbone. The trajectory prediction results of Conv-based models are shown in Table 3.

Similar to the results of LSTM-based models, the ADE and FDE on the pedestrians in front of the ego-vehicle are worse than all pedestrians in Waymo Open Dataset, which indicates the same conclusion that when complicated interactions are involved, getting accurate predictions becomes harder. For data with all pedestrians, the best ADE is

0.327 m achieved by SI-PVI-Conv, while the best FDE is 0.540 m achieved by Social-IWSTCNN. For data with the pedestrians in front of the ego-vehicle, the best ADE is 0.343 m achieved by PEI-Conv as proposed in this paper, which considered the pedestrian and ego-vehicle interaction. The best FDE is 0.578 m, achieved by Social-IWSTCNN. On the Waymo front data, the ADE is improved by 7.0% compared to the Conv model. This shows that the interaction between pedestrians and ego-vehicle influences a pedestrian's behavior, and utilizing this factor can improve the prediction performance.

Table 3. The ADE/FDE metrics (in meters) for non-sequential methods (Conv-based models) consider different factors on the Waymo Open Dataset. The lower, the better. Traj stands for past trajectories of pedestrians. SI stands for social interactions between pedestrians. PVI stands for pedestrian-vehicle interactions. PEI stands for pedestrian and ego-vehicle interaction.

Model	Factors considered for prediction				Waymo (all pedestrians)	Waymo (front pedestrians)
	Traj	SI	PVI	PEI	ADE / FDE	ADE / FDE
Conv	✓				0.334 / 0.571	0.369 / 0.610
Social-IWSTCNN (2021)	✓	✓			0.329 / **0.540**	0.351 / **0.578**
PVI-Conv (2022)	✓		✓		0.346 / 0.586	0.373 / 0.633
SI-PVI-Conv (2022)	✓	✓	✓		**0.327** / 0.543	0.380 / 0.645
PEI-Conv (ours)	✓			✓	-	**0.343** / 0.582
SI-PEI-Conv (ours)	✓	✓		✓	-	0.355 / 0.596
SI-PVI-PEI-Conv (ours)	✓	✓	✓	✓	-	0.369 / 0.612

Compared with the Conv model without any interactions, adding social interactions between pedestrians and adding pedestrian and ego-vehicle interactions can improve the prediction. However, unlike in LSTM-based models, adding pedestrian-vehicle interactions does not improve the prediction accuracy. This could be because for the Conv-based model, extracting the pedestrian-vehicle feature with linear embedding and MLPs and then using convolutional layers may not be the best way to represent the interactions between pedestrians and vehicles. This shows that the most proper feature extraction varies for different prediction models. Although the PEI-Conv and Social-IWSTCNN outperform the Conv model, the combination of the two does not get any further improvement. A better way to combine these features should be further developed in the future.

Compared with the LSTM-based models, the Conv-based models achieve better performance because of the property of the models. The Conv-based networks are better at processing long sequences compared with LSTM-based networks. LSTM-based models treat the “temporal distance” between two time-steps as linear and may “forget” the information of the previous sample in the sequence. In contrast, in Conv-based networks, the convolutional kernels handle the “temporal attention”, and do not have the problem of forgetting previous information. Therefore, the Conv-based models perform better on ADEs. Besides, the LSTM-based models accumulate the errors along with temporal sequences due to the dependency on preceding steps, while the Conv-based models predict all future time-steps simultaneously, so they avoid the error accumulation, thereby getting better FDEs. Therefore, unlike the LSTM-based models that usually get better ADE and FDE simultaneously, the Conv-based models can get the best ADE and FDE from different models.

5. Conclusion and Future Work

For the factors that influence the pedestrian behavior, we have investigated a) the pedestrians’ past trajectories, b) the social interaction between pedestrians, c) the interaction between pedestrians and vehicles, and d) the interaction between pedestrians and the ego-vehicle on both sequential and non-sequential deep learning methods. The results indicate that all these factors can influence the prediction accuracy of pedestrians’ future trajectories. However, we still need to find a better way to aggregate these features to take their advantages. The prediction accuracy of Conv-based models outperforms the LSTM-based models. Future work should focus on developing a proper deep learning network structure to combine these factors. Furthermore, other influencing factors such as the environment could be studied further, and the most influential factors could be used to improve the behavior prediction.

Acknowledgements

This research is funded by the European research project "SHAPE-IT - Supporting the Interaction of Humans and Automated Vehicles: Preparing for the Environment of Tomorrow" (cf. https://www.shape-it.eu/), that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement 860410. The authors would like to thank Asymptotic AI (cf. https://www.asymptotic.ai/) and CGit (cf. https://www.cgit.se/) for providing the computational resource and the GPU platform AiQu.

References

Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Li, F.F., Savarese, S., 2016. Social lstm: Human trajectory prediction in crowded spaces, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 961–971.
Eiffert, S., Li, K., Shan, M., Worrall, S., Sukkarieh, S., Nebot, E., 2020. Probabilistic crowd gan: Multimodal pedestrian trajectory prediction using a graph vehicle-pedestrian attention network. IEEE Robotics and Automation Letters 5, 5026–5033.
Fernando, T., Denman, S., Sridharan, S., Fookes, C., 2018. Soft+ hardwired attention: An lstm framework for human trajectory prediction and abnormal event detection. Neural networks 108, 466–478.
Giuliari, F., Hasan, I., Cristani, M., Galasso, F., 2021. Transformer networks for trajectory forecasting, in: 2020 25th International Conference on Pattern Recognition (ICPR), IEEE. pp. 10335–10342.
Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., Alahi, A., 2018. Social gan: Socially acceptable trajectories with generative adversarial networks, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2255–2264.
Kotseruba, I., Rasouli, A., Tsotsos, J.K., 2021. Benchmark for evaluating pedestrian action prediction, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 1258–1268.
Lerner, A., Chrysanthou, Y., Lischinski, D., 2007. Crowds by example, in: Computer graphics forum, Wiley Online Library. pp. 655–664.
Mohamed, A., Qian, K., Elhoseiny, M., Claudel, C., 2020. Social-stgcnn: A social spatio-temporal graph convolutional neural network for human trajectory prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 14424–14432.
Nikhil, N., Tran Morris, B., 2018. Convolutional neural network for trajectory prediction, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops.
Ohn-Bar, E., Trivedi, M.M., 2016. Looking at humans in the age of self-driving and highly automated vehicles. IEEE Transactions on Intelligent Vehicles 1, 90–104.
Pellegrini, S., Ess, A., Schindler, K., Van Gool, L., 2009. You'll never walk alone: Modeling social behavior for multi-target tracking, in: 2009 IEEE 12th International Conference on Computer Vision, IEEE. pp. 261–268.
Rasouli, A., Tsotsos, J.K., 2019. Autonomous vehicles that interact with pedestrians: A survey of theory and practice. IEEE Transactions on Intelligent Transportation Systems 21, 900–918.
Sadeghian, A., Kosaraju, V., Sadeghian, A., Hirose, N., Rezatofighi, H., Savarese, S., 2019. Sophie: An attentive gan for predicting paths compliant to social and physical constraints, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
Schneider, N., Gavrila, D.M., 2013. Pedestrian path prediction with recursive bayesian filters: A comparative study, in: German Conference on Pattern Recognition, Springer. pp. 174–183.
Shirazi, M.S., Morris, B., 2015. Observing behaviors at intersections: A review of recent studies & developments, in: 2015 IEEE Intelligent Vehicles Symposium (IV), IEEE. pp. 1258–1263.
Song, X., Chen, K., Li, X., Sun, J., Hou, B., Cui, Y., Zhang, B., Xiong, G., Wang, Z., 2020. Pedestrian trajectory prediction based on deep convolutional lstm network. IEEE Transactions on Intelligent Transportation Systems 22, 3285–3302.
Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B., et al., 2020. Scalability in perception for autonomous driving: Waymo open dataset, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2446–2454.
WHO, 2018. Global status report on road safety 2018: Summary. Report. World Health Organization.
Yang, B., Zhan, W., Wang, P., Chan, C., Cai, Y., Wang, N., 2021. Crossing or not? context-based recognition of pedestrian crossing intention in the urban environment. IEEE Transactions on Intelligent Transportation Systems .
Zhang, C., Berger, C., . Learning the pedestrian-vehicle interaction for pedestrian trajectory prediction, in: 2022 the 8th International Conference on Control, Automation and Robotics (ICCAR), IEEE. pp. 230–236.
Zhang, C., Berger, C., Dozza, M., 2021. Social-iwstcnn: A social interaction-weighted spatio-temporal convolutional neural network for pedestrian trajectory prediction in urban traffic scenarios, in: 2021 IEEE Intelligent Vehicles Symposium (IV), IEEE. pp. 1515–1522.
Zhang, H., Liu, Y., Wang, C., Fu, R., Sun, Q., Li, Z., 2020. Research on a pedestrian crossing intention recognition model based on natural observation data. Sensors 20, 1776.
Zhang, P., Ouyang, W., Zhang, P., Xue, J., Zheng, N., 2019. Sr-lstm: State refinement for lstm towards pedestrian trajectory prediction, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
Zhou, W., Berrio, J.S., De Alvis, C., Shan, M., Worrall, S., Ward, J., Nebot, E., 2020. Developing and testing robust autonomy: The university of sydney campus data set. IEEE Intelligent Transportation Systems Magazine 12, 23–40.